THE UNIVERSITY OF
WARWICK

Original citation:
Morris, Kyle L., Rodger, Alison, Hicks, Matthew R., Debulpaep, Maya, Schymkowitz, Joost, Rousseau, Frederic and Serpell, Louise C.. (2013) Exploring the sequence–structure relationship for amyloid peptides. Biochemical Journal, Vol.450 (No.2). pp. 275-283.
Permanent WRAP url:
http://wrap.warwick.ac.uk/53122

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes the work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0) license and may be reused according to the conditions of the license. For more details see: http://creativecommons.org/licenses/by-nc-nd/3.0/

A note on versions:
The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwickpublicationswrap
highlight your research

http://go.warwick.ac.uk/lib-publications

BJ Structure

View this PDF in Utopia Documents for more interactive functionality
Download FREE software now

Biochem. J. (2013) **450**, 275–283 (Printed in Great Britain) doi:10.1042/BJ20121773

online data

Exploring the sequence–structure relationship for amyloid peptides

Kyle L. MORRIS*, Alison RODGER†, Matthew R. HICKS‡, Maya DEBULPAEP§, Joost SCHYMKOWITZ‖, Frederic ROUSSEAU‖ and Louise C. SERPELL*

*School of Life Sciences, University of Sussex, Falmer, East Sussex BN1 9QG, U.K., †Department of Chemistry, University of Warwick, Coventry CV4 7AL, U.K., ‡School of Biosciences, University of Birmingham, Birmingham B15 2TT, U.K., §VIB Switch Laboratory, Free University of Brussels, Pleinlaan 2, Brussels 1050, Belgium, and ‖VIB Switch Laboratory, Department of Cellular and Molecular Medicine, University of Leuven, Leuven 3000, Belgium

Amyloid fibril formation is associated with misfolding diseases, as well as fulfilling a functional role. The cross-β molecular architecture has been reported in increasing numbers of amyloid-like fibrillar systems. The Waltz algorithm is able to predict ordered self-assembly of amyloidogenic peptides by taking into account the residue type and position. This algorithm has expanded the amyloid sequence space, and in the present study we characterize the structures of amyloid-like fibrils formed by three peptides identified by Waltz that form fibrils but not crystals. The structural challenge is met by combining electron microscopy, linear dichroism, CD and X-ray fibre diffraction. We propose structures that reveal a cross-β conformation with 'steric-zipper' features, giving insights into the role for side chains in peptide packing and stability within fibrils. The amenity of these peptides to structural characterization makes them compelling model systems to use for understanding the relationship between sequence, self-assembly, stability and structure of amyloid fibrils.

Key words: amyloid, amyloidogenic peptide, Waltz algorithm, X-ray fibre diffraction.

INTRODUCTION

Amyloid is a fibrillar self-assembled non-covalent proteinacious polymer [1]. A great number of proteins have been shown to be able to access a self-assembled β-sheet conformation and it has been suggested that under the right conditions any protein may form amyloid. It has been proposed that the amyloid conformation represents an energetically stable conformation that may be accessed by any polypeptide [2]. Indeed, a diverse range of proteins form amyloid with large variations in protein sequence and amyloidogenic precursor structure. Proteins that form amyloid are typically associated with a degenerative disease; amyloid-β with Alzheimer's disease, transthyretin with familial amyloidotic polyneuropathy, IAPP (islet amyloid polypeptide) with Type 2 diabetes and prion or PrP (prion protein) with TSEs (transmissible spongiform encephalopathies). Potentially other diseases that have previously had unknown underlying pathology are caused by an aggregating peptidic monomer. For example, it has recently been suggested that phenylketonuria may be caused by an accumulation of amyloid-like assemblies of the single amino acid phenylalanine [3]. The ability of a single amino acid to self-assemble in an ordered manner, when compared with the 42-residue protein amyloid-β [4] or the larger ~200 residue PrP [5], is a striking example of the variety found in amyloid systems. This variety is not only fascinating, but promises to have great utility in the application of such systems to bionanomaterial design.

The number of amyloidogenic systems associated with disease has led to numerous investigations into their fibrillar structure [4,6–8], each revealing a common shared architecture named cross-β [9]. The cross-β architecture is defined by β-strands orientated perpendicular to the fibril long axis stacked to form long β-sheets that extend into a fibril. Two or more of these sheets may associate laterally to form the protofilaments of mature amyloid fibrils. These molecular details were first proposed on the basis of the interpretation of X-ray fibre diffraction data from amyloid-like systems [8,10,11]. This architecture has since been found in various other amyloid systems using ssNMR (solid-state NMR) [4,6,7] and also X-ray crystallography [12]. These same advances have further provided structural details that have revealed insights into side-chain interactions that form the basis for the β-sheet association [6,8,12].

Lately, efforts have been made to expand the sequence space of structurally characterized amyloid systems. These sequences may be identified within larger amyloidogenic sequences or be identified by bioinformatics. Identification of amyloidogenic sequences via a bioinformatics approach is not only able to expand the amyloid sequence space, but in the identification algorithms required, highlight the characteristics of a sequence that give it a propensity to assemble [13]. These algorithms are numerous and based on varying principles, but typically identify short peptide sequences, often hexapeptides [14–18]. A systematic study of the peptide with the sequence STVIIE revealed the important positional and chemical characteristic of a hexapeptide by studying the amyloid-forming propensity of variants of the peptide sequence [19]. We previously reported a new algorithm, called Waltz, for identifying short amyloidogenic sequences [18] on the basis of a sequence position scoring matrix. It has successfully identified a small library of amyloidogenic peptide sequences that self-assemble in solution to form ordered amyloid aggregates.

Charge, hydrophobicity and aromaticity have all been implicated in the structural basis for assembly [20]. The present study provides a basis from which to understand the relationship between sequence and the ability of the peptide to form amyloid fibrils and their resulting structures. We report on the structural and biophysical characterization of three of the peptides, identified by Waltz, able to form highly ordered amyloid-like fibrils with the sequences HYFNIF, RVFNIM and VIYKI. Waltz was designed to identify not only amyloid-forming sequences, but also to

Abbreviations used: CCD, charge-coupled-device; FTIR, Fourier-transform infrared; LD, linear dichroism; PrP, prion protein; TEM, transmission electron microscopy; XRFD, X-ray fibre diffraction.

[1] To whom correspondence should be addressed (email l.c.serpell@sussex.ac.uk).

© The Authors Journal compilation © 2013 Biochemical Society

Biochemical Journal www.biochemj.org

better distinguish between amyloid sequences and amorphous aggregates [18]. The characterization of these three Waltz peptides further re-enforces the algorithm's ability to do this. This investigation and the analysis of fibrillar systems uses time-resolved XRFD (X-ray fibre diffraction) and analysis of LD (linear dichroism) and CD spectroscopy to identify structural characteristics and produce models representative of the fibrillar semi-crystalline structure. In doing so we reveal insights into the relationship between sequence and structure relevant to understanding amyloid formation in disease and as a basis for designing new self-assembling bionanomaterials.

EXPERIMENTAL

Peptide synthesis and materials

The short peptides were synthesized as freeze-dried powder with N-acetylated and C-amidated capped termini to >95 % purity by HPLC (JPT Peptide Technologies). Peptides were incubated in water at 10 mg/ml under quiescent conditions at room temperature (20–25 °C) for 1 week unless otherwise stated. Fibres allowed to assemble for longer incubation periods of up to 3 months are referred to as mature. Water was purified by reverse osmosis and filter-sterilized using 0.2 μm membranes (Minisart).

TEM (transmission electron microscopy)

Samples were examined by TEM to confirm fibril formation. Samples were diluted to 0.25–1 mg/ml prior to TEM visualization. For grid preparation, a 4 μl drop of fibril solution was incubated on Formvar/Carbon 400 Mesh Copper grids (Agar Scientific) for 1 min, followed by a 1 min wash with 4 μl of water and negatively stained twice with 4 μl of 2 % uranyl acetate for 1 min. Between each stage, excess liquid was removed by blotting with filter paper. Care was taken not to dry out the grid until after the final negative stain step. Grids were visualized using a Hitachi-7100 transmission electron microscope running at 100 kV. Images were taken using an axially mounted (2000 pixels × 2000 pixels) Gatan Ultrascan 1000 CCD (charge-coupled-device) camera. The program ImageJ (http://rsbweb.nih.gov/ij/) was used to analyse TEM images [21] and individual fibril morphologies were inspected after applying the in-built fast Fourier-transform band-pass filter.

CD

A Jasco J-715 spectropolarimeter with a peltier temperature control system was used to collect CD spectra. All measurements were collected at 20 °C with a sample concentration of 1–2 mg/ml, 180–320 nm with a pitch of 0.1 nm at a scan speed of 50 nm/min, a response time of 4 s, slit widths of 1 nm and standard sensitivity. Control buffer spectra were averaged from triplicate accumulations and subtracted from the averaged triplicate accumulations of samples. Pathlengths were adjusted according to the degree of signal against PMT (photomultiplier tube) high-tension voltage to ensure high data quality and were 0.1 mm using quartz demountable cells (Starna Scientific). To check for contributions arising from LD artefacts due to orientation effects, cuvettes were rotated by 90° in the instrument. Where artefacts arising from orientation of fibrils were found, a combination of tip (Sonics Vibra-Cell VCX500, 20 kHz, 40 % amplitude, 2 min) and water bath (Fisher Scientific FB15051, 37 kHz, 1 min) sonication was used to disrupt alignment.

Secondary structure analysis was performed on spectra confirmed to contain only CD signals using the online server Dichroweb [23] with the CDSSTR analysis programme [24] and best available reference set: SP175 190–240 nm [25]. Careful attention was paid to the closeness of fit between secondary structure model predictions and the experimental data where low NRMSD (normalized root mean square deviation) values indicate a high goodness of fit of the secondary structure model prediction and successful analysis [26].

LD

A Jasco J-815 spectropolarimeter modified for LD was used to collect LD spectra. Measurements were taken in a 0.5 mm pathlength at 20 °C at 200 μg/ml, 180–320 nm with a pitch of 0.2 nm at a scan speed of 100 nm/min, a response time of 1 s, bandwidth of 1 nm and standard sensitivity. Spectra were collected as an average of three accumulated measurements and processed by subtraction of the background signal from water and zeroing at 300–320 nm. Three channels were monitored; LD, HT[V] (high tension [voltage]) and absorbance. To unambiguously detect LD signals, fibrillar samples were aligned by a Couette flow cell apparatus [27] using rotation speeds of 1500 rev./min, creating fibre alignment parallel with the orientation axis (//). Assembly of the Couette flow cell apparatus was found to align fibres perpendicular to the orientation axis ($\perp$) and thus comparison spectra were also collected with no Couette flow.

FTIR (Fourier-transform infrared) spectroscopy

Infrared spectra were recorded using a Bruker Tensor 27 infrared spectrophotometer equipped with a Bio-ATR II accessory. Spectra were recorded of dried films at a spectral resolution of 4 cm^{-1} with 120 accumulations performed per measurement at a wavenumber range of 900–3500 cm^{-1}. Buffer and baseline subtraction from the obtained spectra were made with rescaling in the wave number range of 900–1800 cm^{-1}.

XRFD analysis

Fibril samples were aligned by a variety of methods to produce different textures. Fibrous-textured alignments were created by suspending a 10 μl droplet of 10 mg/ml fibril solution between two wax-tipped 1.2 mm outside diameter, 0.94 mm internal diameter borosilicate capillaries (Harvard apparatus) and placing in a parafilm-sealed Petri dish to air-dry at room temperature. Film-textured alignments were formed by drawing a 10–50 μl droplet of 10 mg/ml fibril solution into a X-ray transmissible 0.7 mm borosilicate capillary (Capillary Tube Suppliers) and sealed at one end to prevent further capillary action. This solution was allowed to dry by evaporation to create a film-texture. These methods are further described in [28]. In order to monitor real-time fibre alignment, a cell similar to that described previously was used [29]. Briefly, the cell comprises a semi-enclosed chamber that can support the wax-tipped capillaries used for alignment, such that this process can be monitored by X-ray fibre diffraction in real-time. Samples for fibre diffraction were tested on a home source Rigaku 007HF Cu Kα [$\lambda = 1.5419$ Å (1 Å = 0.1 nm)] rotating anode generator with VariMax-HF mirrors and Saturn 944 + CCD detector. Additionally, synchrotron data were collected at the Diamond I24 microfocus beamline with a wavelength of 0.9778 Å and an MARCCD detector. Clearer [30] was used to measure signal positions in XRFD patterns. Graphical traces were produced by sampling a 60° radial slice of the

© The Authors Journal compilation © 2013 Biochemical Society

Table 1 The Waltz peptide sequences and their origin with regard to the full-length proteins in which they are found

The Waltz algorithm predicts the likelihood of an amino acid being found at a particular position of an amyloidogenic sequence. The probability of a residue occurring at a particular position in an amyloidogenic peptide as predicted by Waltz is expressed as a log-odd score [18]. See also Supplementary Figure S1 (at http://www.biochemj.org/bj/450/bj4500275add.htm). eIF, eukaryotic initiation factor.

	Origin						Residue position scoring					
Waltz sequence	UniProt ID	Length	Segment	Name	Function	Organism	1	2	3	4	5	6
HYFNIF	P54132	1417	27–33	Bloom syndrome protein	DNA helicase	Human	0	0	1	1	2	0
RVFNIM	P20042	333	182–187	eIF-2	Translation regulation	Human	−1	0	1	1	2	0
VIYKI	P07182	284	183–187	Chorion protein	Egg shell protein	*Drosophila*	0	1	−1	−2	−2	

XRFD on either meridional or equatorial axes and plotted using Braggs law. Determined diffraction signals were entered into the unit cell determination module within Clearer [30] and possible unit cells were explored by comparing experimental diffraction signal positions with calculated signals and examining for logical indexing schemes.

Modelling

Initial peptide models were built in Insight II (Molecular Dimensions) in an ideal β-strand conformation. Rotamer conformations were assigned using the Dunbrack backbone-dependent rotamer library in PyMOL (http://www.pymol.org), variable architectures were explored and assessed using Molprobity [31], checking for clashes, favourable Φ-Ψ angles and side-chain positions. Crystal lattices were constructed in PyMOL using the unit cell determined by Clearer [30], and these were minimized to remove side-chain clashing and unfavourable conformations. Explicit lattice minimizations were performed using the CHARMM Prot_all_22_forcefield [33] through NAMD [34] as implemented in VMD [35] in a (3 3 3) solvated lattice.

Simulated XRFD and model validation

Using Clearer, fibre-diffraction patterns were simulated from minimized models. Briefly, the central model of a minimized lattice was used to construct a fibre texture from which the intensity and position of reflections were simulated. The diffraction settings were equivalent to the experimental collection parameters (i.e. specimen-to-detector distance, wavelength and detector parameters). The fibre disorder parameters σ_θ and σ_Φ were 0.2 and ∞ respectively, and the crystallite size was set to 400 Å^3 with a sampling interval of 1 pixel. All other simulation settings were default.

Pattern comparison was initially qualitative and based on visual comparison of simulated XRFD patterns against the experimental pattern. To more accurately identify closeness of fit (RF), the experimental and simulated reflections were tabulated and a quantitative comparison was made on the basis of signal position and relative intensity.

RESULTS

Sequence identification and origin

The three short peptides, identified by Waltz and shown in the present paper, are five or six residues in length and taken from much larger polypeptides with variable native functions, as shown in Table 1. Of the three sequences, the crystal structure of HYFNIF within the context of the full-length native protein structure is available for assessment (PDB code 2KV2). Interestingly the predicted secondary structure by PSIPred [36] and observed secondary structure are different (Supplementary Figure S1 at http://www.biochemj.org/bj/450/bj4500275add.htm). Although HYFNIF was identified by Waltz and predicted to form β-strands, the crystal structure of the native protein reveals this sequence adopts an α-helical structure. RVFNIM and VIYKI are predicted to respectively adopt α-helical and β-strand conformations.

In the case of the hexapeptides, the terminal residues are either neutral or slightly unfavoured in the Waltz scoring algorithm, indicating that the core residues are key and may represent a possible core interaction motif. This core sequence is Phe-Asn-Ile at positions 3–5 in the hexapeptides, whereas in the case of the pentapeptide the equivalent positions 3–5 have the sequence Tyr-Lys-Ile. In general the core may be described as having the aromatic pattern A-X-X and the hydrophobicity pattern H-X-H.

Morphology characterization by TEM

TEM of the Waltz peptides revealed that in water they spontaneously form amyloid-like fibrillar morphologies, with widths of approximately 20 nm (Figure 1). Each peptide fibril exhibits characteristic morphologies, inferring the complex association of individual protofilaments in a hierarchical structure.

Fibrils formed by HYFNIF exhibited the greatest range of morphologies, which can broadly be described as having a twisted morphology. The paired filament helices (Figure 1A, i) have a greatest width of 17.04 nm (S.D. $\pm$ 1.06; $n = 7$) and periodicity of 95.16 nm (S.D. $\pm$ 3.15; $n = 6$). Other helical arrangements of ribbons are observed with variable widths and periodicities (Figure 1A, ii and iii). Following long incubation (several months) the fibrils developed into structures that may represent a tubular architecture with widths of 26.03 nm (S.D. $\pm$ 3.29; $n = 3$) (Figure 1A, iv and Supplementary Figure S2 at http://www.biochemj.org/bj/450/bj4500275add.htm).

RVFNIM was more consistently found to form tightly wound 'ropes' with a width of 17.61 nm (S.D. $\pm$ 1.23; $n = 5$) with some indication of protofilament structure (Figure 1B, i and ii). Some examples of rope-like protofilaments twisting into a helical arrangement were also observed (Figure 1B, iii) with widths at their widest points of 19.69 nm (S.D. $\pm$ 1.59; $n = 28$). The helical pitch of these morphologies varies dramatically from 147.81 nm (S.D. $\pm$ 21.40; $n = 11$) to indeterminate over a single TEM micrograph (Figure 1B, iv).

VIYKI exhibited the most regular morphology where all observed fibres formed twisted 'ropes' (Figure 1C, i–iv) with a width of 21.0 nm (S.D. $\pm$ 1.2; $n = 6$) in a helical twist with a periodicity of 58.1 nm (S.D. $\pm$ 1.7; $n = 6$).

Secondary structure determination by CD and FTIR

The secondary structures of the fibrillar Waltz peptides were investigated using CD and FTIR spectroscopy. Characterization

© The Authors Journal compilation © 2013 Biochemical Society

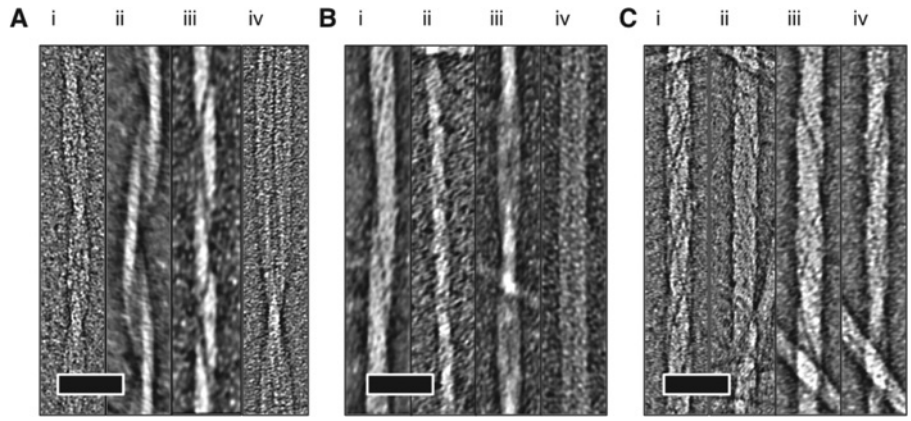

Figure 1 The characteristic morphologies of Waltz fibrils

(**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI as analysed by TEM. Micrographs have been cropped and processed by fast Fourier-transform bandpass filtering in ImageJ [21]. The scale bars represent 50 nm. See also Supplementary Figure S2 (at http://www.biochemj.org/bj/450/bj4500275add.htm).

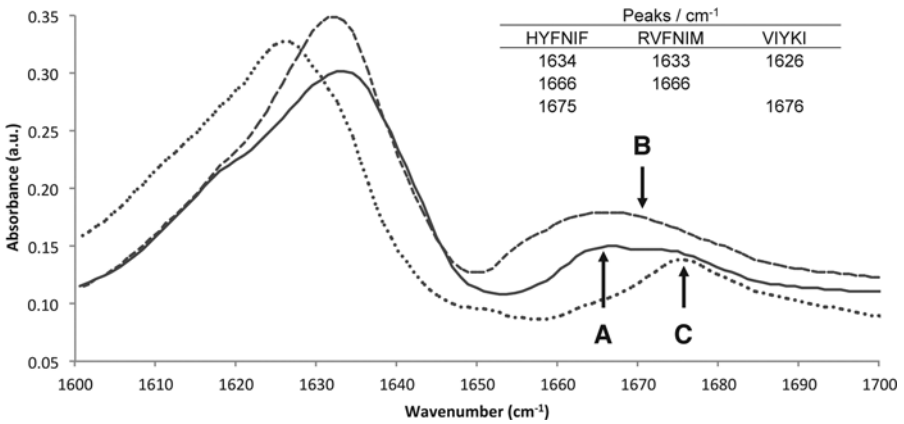

Figure 2 The secondary structure of the Waltz fibrils

The amide I FTIR region of dried films of mature assembled (**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI, each indicating the presence of a β-sheet conformation. See also Supplementary Figure S3 (at http://www.biochemj.org/bj/450/bj4500275add.htm).

by CD is complicated by both the anisotropic nature and high aromatic content of these systems. A classic β-sheet spectrum typically has a positive and a negative maximum at $\sim$195 and $\sim$216 nm respectively. The CD spectra from the Waltz peptides are more complicated. Each CD spectrum exhibits a strong dependence on sample orientation, with signal intensities that are far greater than normal (Supplementary Figure S3 at http://www.biochemj.org/bj/450/bj4500275add.htm) and signal positions that are not as expected for typical β-sheet peptides. The fibril structure and thus potential for the shear alignment of these systems may give rise to LD artefacts in the CD measurements. To obtain CD spectra with minimal contribution from artefactual LD signals, the Waltz fibrils were sonicated in an attempt to abolish effects from alignment (Supplementary Figure S3) as reported elsewhere [22]. TEM confirmed that, following sonication, fibrils remained present in the solution, but sample orientation CD signal dependence was no longer observed, revealing the true CD signals representative of the Waltz peptides. Dichroweb analysis revealed predominantly β-sheet and other non-helical structures (Supplementary Table S1 at http://www.biochemj.org/bj/450/bj4500275add.htm). We presume the non-helical conformations represent peptides in a random-coil conformation that have been liberated from fibres, but the remaining predominant β-strand structures constitute the Waltz fibrils.

Since the high aromatic content and anisotropy complicates CD analysis, we corroborated these results with FTIR of mature Waltz fibrils. The expected absorption bands for β-sheet structure with two maxima at 1626–1633 and 1666–1676 cm^{-1} were observed (α-helices exhibit bands at $\sim$1654 cm^{-1} [37]) for all three peptide assemblies (Figure 2). FTIR has been used to distinguish between parallel and antiparallel β-sheets by the presence of the longer wavenumber band at 1695 cm^{-1} [38], but the reliability of this interpretation is still a matter of debate [37] and so here we interpret the FTIR only in terms of the presence of β-sheet conformation.

Chromophore orientation determination by LD

Biophysical analyses identified the β-sheet-rich structure of the Waltz peptide fibrils. The presence of LD artefacts in the CD spectra suggests that a regular and ordered arrangement of structural elements is apparent in the assemblies. The alignment of the Waltz peptide fibrils in the initial CD experiments was the result of sample loading. This cannot be controlled and so we chose to measure the LD signals using a micro-volume Couette

© The Authors Journal compilation © 2013 Biochemical Society

flow cell apparatus [39] as this provides a reproducible alignment methodology to gain more information regarding the orientation of structural elements and chromophores within the amyloid fibrils.

LD on fibrillar systems, in particular amyloid, has been reported previously [22,27,40,41]. Transition moments in chromophores that are regularly orientated produce positive and negative LD signals depending on their orientation. The orientation of chromophores relative to the fibre axis can thus be determined from the magnitude and sign of their LD signals [42]. The LD spectra for the Waltz fibrils were collected when aligned parallel with (//) and perpendicular to (⊥) the orientation axis (Figure 3), where in our instrument the orientation axis is defined as horizontal. Our intention had been to observe orientation under Couette flow; however, the signals shown in Figure 3 indicate that the loading of the sample provided sufficient shear force to induce perpendicular fibril orientation. All three peptides show a large negative signal at ~200 nm arising from the π-π* transition of the β-sheet peptide backbone. The sign of the β-sheet π-π* transition is thus consistent with its polarization being parallel with the fibre long axis. Since this transition occurs perpendicular to the β-strand direction [42], the data are consistent with β-strands arranged perpendicular to the fibre long axis and are consistent with the cross-β architecture, as observed for other self-assembled systems [22,40].

Couette flow created significantly more orientation for HYFNIF fibrils, not enough orientation for RVFNIM fibrils to overrule the loading-induced orientation and some orientation for VIYKI fibrils. For HYFNIF and VIYKI fibrils, the signals were also observed in the region expected to arise from aromatic LD spectra (Figure 3). RVFNIM fibrils gave no evidence of orientation of its aromatic chromophores, although this is likely to be due to the fact that phenylalanine has a low absorption coefficient. By way of contrast, HYFNIF and VIYKI have clearly resolved aromatic tyrosine LD signals, the L_a (230 nm) and L_b (275 nm) transitions, indicating a regular ordered arrangement of these aromatic residues in these structures. The splitting of the L_b tyrosine transition into two components is an excitonic coupling effect indicating that the tyrosine residues in these systems are closely associated and possibly involved in π-π stacking interactions [43]. This phenomenon has been previously noted in LD spectra from fibrils formed by the heptapeptide GNNQQNY [40]. Thus the LD data indicate that the tyrosine long axes for both HYFNIF and VIYKI fibrils are orientated greater than 54.7° from the fibre axis, whereas the tyrosine short axes are orientated less than 54.7° and may adopt a stacked geometry. Knowing the β-sheet content and orientation with the likelihood of aromatic stacking interactions we sought to gather more structural details from XRFD.

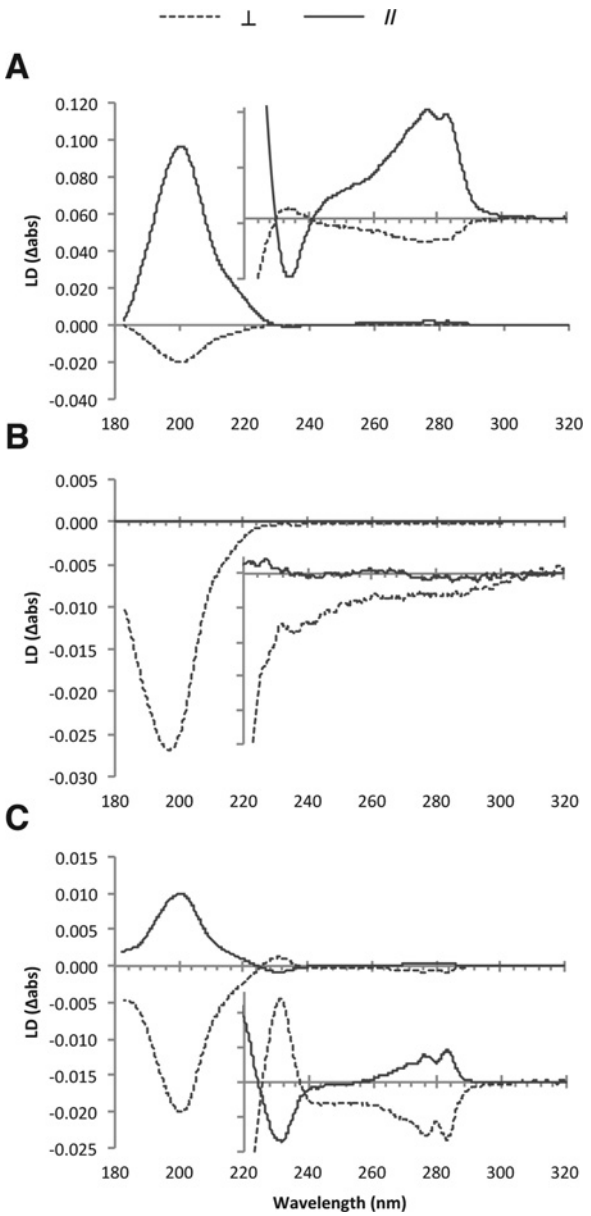

Figure 3 The LD spectra arising from the Waltz fibrils

Alignment of fibrils was induced by shear force perpendicular (⊥; broken lines) and Couette flow parallel (//; continuous lines) to the orientation axis of (**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI. Peaks are tabulated in Supplementary Table S2 (at http://www.biochemj.org/bj/450/bj4500275add.htm).

XRFD from Waltz peptide fibrils

Fibrils of Waltz assemblies were readily aligned, producing fibrous-textured alignments, by methods previously described [28]. These alignments exhibit birefringence by cross-polarizing light microscopy, indicating a para-crystalline order of the assemblies (Supplementary Figure S4 at http://www.biochemj.org/bj/450/bj4500275add.htm). Each system exhibited the major meridional and equatorial features associated with the cross-β architecture commonly observed for amyloid-like assemblies [9]. These results are thus consistent with the spectroscopic observations and the deduction of the cross-β arrangement within these peptide assembly systems. The reflections arising from the repetitive interatomic separations within the Waltz fibrils differ between the systems, indicating differences in the packing arrangements of the different peptides.

We also monitored XRFD in real-time over the course of fibre alignment. To our knowledge, we report for the first time the XRFD exhibited by an amyloid-like system over the course of fibre alignment in the hydrated, semi-hydrated and dried aligned state as shown in Figure 4. Importantly, we find that the reflection positions arising from the amyloid architecture are present, and the diffuse scattering from water (~3.5 Å) is still observable. This supports the view that the structures of the fibres are the same in the solution and dried state. The same phenomenon was confirmed for each Waltz system individually (Supplementary Figures S4D and S4E). We also explored the effect of texture on the observed XRFD patterns. Samples of

© The Authors Journal compilation © 2013 Biochemical Society

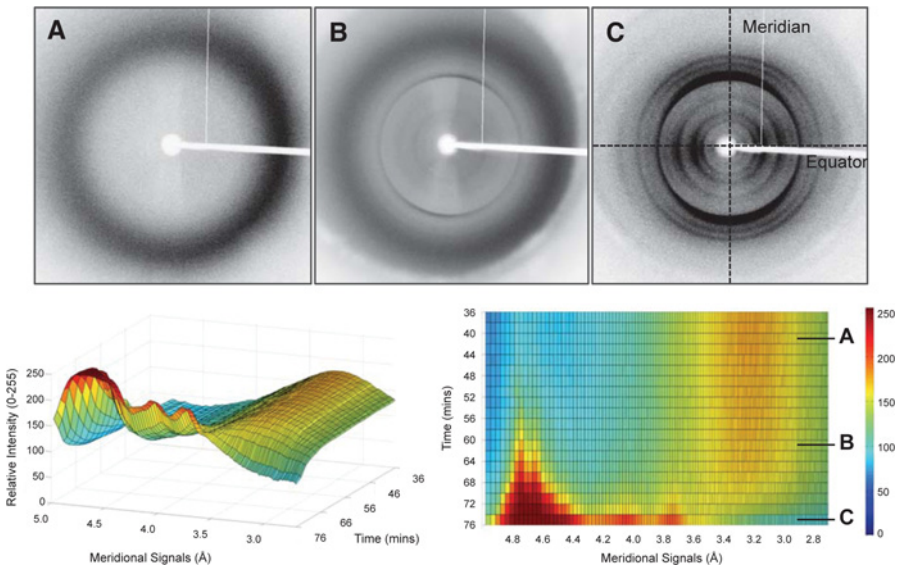

Figure 4 Real-time XRFD of VIYKI over the course of alignment

In the (**A**) hydrated, (**B**) semi-hydrated and (**C**) dried and aligned state. The corresponding graphical representation shows that the meridional signals arising from the cross-β structure are present before the complete loss of water and thus are present in the normal hydrated state. All fibre axes are vertical. See also Supplementary Figures S4(D) and S4(E) (at http://www.biochemj.org/bj/450/bj4500275add.htm).

fibres were aligned to produce film-textured alignments where all fibre axes are aligned parallel with the film plane. We note that the film-textured alignments also report the same diffraction signals as the fibrous alignments (Supplementary Figure S5 at http://www.biochemj.org/bj/450/bj4500275add.htm), but believe that the film texture may abolish contribution from fibril packing. In the case of RVFNIM, additional equatorial information is observed in the film-textured alignment and so this pattern was used in subsequent analysis.

The relatively large number of reflections exhibited by these systems is indicative of a para-crystalline order greater than that typically observed for amyloid-like systems [9]. The quality of the patterns thus presents an opportunity to determine a unit cell on the basis of the reflection positions (Supplementary Table S3 at http://www.biochemj.org/bj/450/bj4500275add.htm) and model the molecular structure within the Waltz fibrillar assemblies. All three peptides share a strong meridional reflection, the distance of which varies between 4.66 and 4.76 Å (Supplementary Table S3), arising from the separation between hydrogen-bonded β-strands along the fibre axis. Higher-order reflections of the principal meridional reflections are present, but the precise position of the meridional reflections is dependent on β-strand separation and relative displacement of these along the fibre axis. Knowing the fibre axis repeat, the equatorial reflections were indexed and predicted unit cell dimensions were determined using Clearer [30]. The determined unit cell dimensions for each system (Supplementary Table S4 at http://www.biochemj.org/bj/450/bj4500275add.htm) are defined as: a, the β-strand chain length; b, the β-sheet spacing; and c, the hydrogen bonding spacing along the fibre axis.

The strongest equatorial reflections pertaining to the distance between β-sheets for the peptides HYFNIF, RVFNIM and VIYKI are 12.4, 10.8 and 9.21 Å respectively. This information allowed for the basic arrangement of the peptides within the repeating cells to be constructed. From X-ray crystallography analysis of amyloidogenic short peptides, a range of cross-β architectures have been proposed and grouped into eight classes [12]. The arrangement of β-strands within β-sheets can be antiparallel, β-sheet relative orientation (face-to-face or face-to-back) and β-sheet direction (up or down). This information was used as the basis for using molecular modelling to explore the possible arrangements of the Waltz peptides within their respective repeating cells (Supplementary Online Data at http://www.biochemj.org/bj/450/bj4500275add.htm).

Modelling and assessment by simulated XRFD

Diffraction patterns were simulated from models systematically constructed (Supplementary Figure S6 at http://www.biochemj.org/bj/450/bj4500275add.htm) using the descriptions of parallel amyloid structural classes [12] and quantitatively compared with the experimental diffraction patterns (Supplementary Table S5 at http://www.biochemj.org/bj/450/bj4500275add.htm). The model structures shown in Figure 5 are those that were explored and found to best fit the diffraction data. A comparison between the simulated and experimental diffraction patterns confirming the validity of the constructed models is shown in Figure 6.

Figure 5 shows the most representative models for each of the Waltz systems. Of the models constructed, the majority of simulated patterns compared well with experimental fibre diffraction data, and so we postulate that the experimental fibre diffraction patterns may result from a mixture of architectures. However, the structures shown in Figure 5 represent the predominant representative architectures. The modelled β-sheet separations closely correspond to the major equatorial reflections (Supplementary Table S3) and the β-strand separation along the fibre axis varies between 4.70 and 4.78 Å (Supplementary Table S4). The low-resolution reflections found on the HYFNIF and VIYKI patterns were not easily reproduced, but we attribute this to our simulations modelling the structures as a continuous lattice rather than discrete protofilaments. HYFNIF is found to adopt an up-up architecture where two sheets are displaced by $b/2$, where the β-sheets are parallel with respect

© The Authors Journal compilation © 2013 Biochemical Society

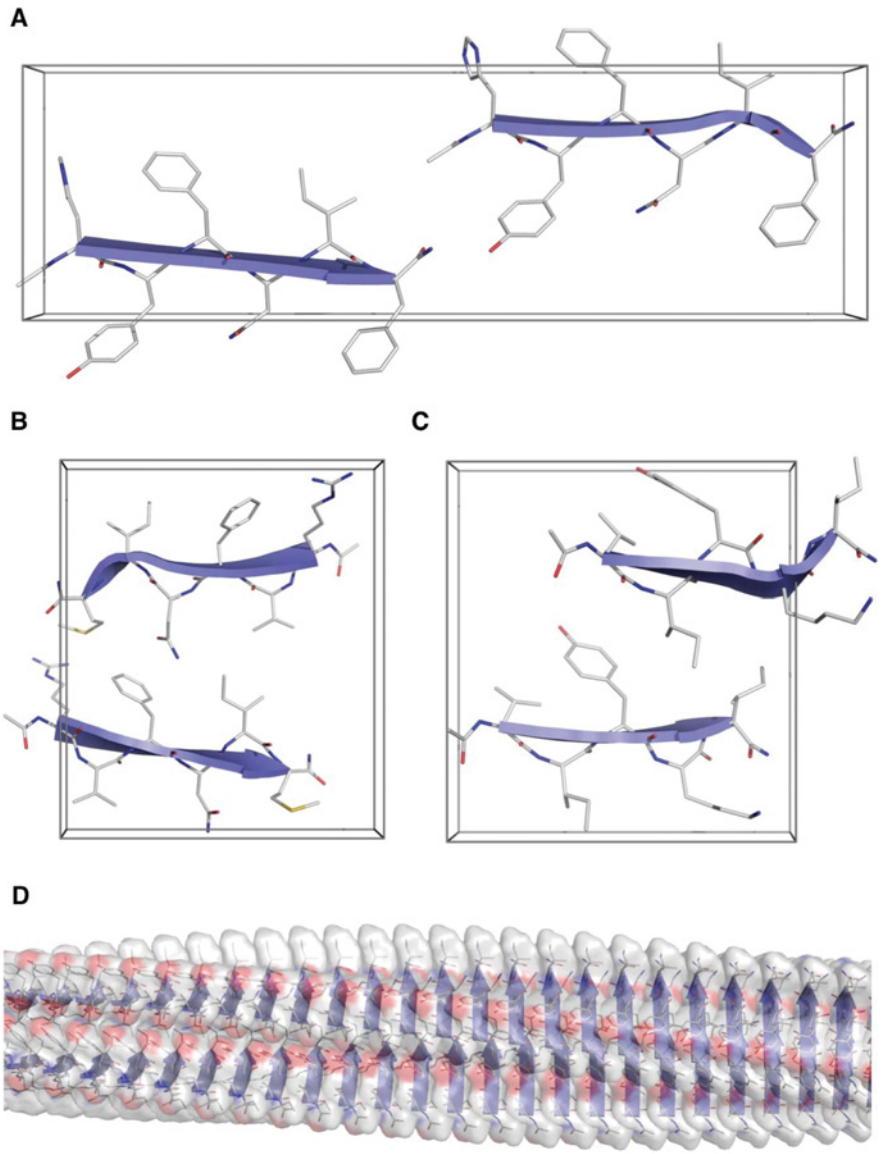

Figure 5 Representative models of the Waltz peptides

The peptides are shown modelled into the unit cells determined from XRFD: (**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI. The visualization (**D**) shows the possible protofilament context of VIYKI. See also Supplementary Figure S6 and the Supplementary Online Data (at http://www.biochemj.org/bj/450/bj4500275add.htm).

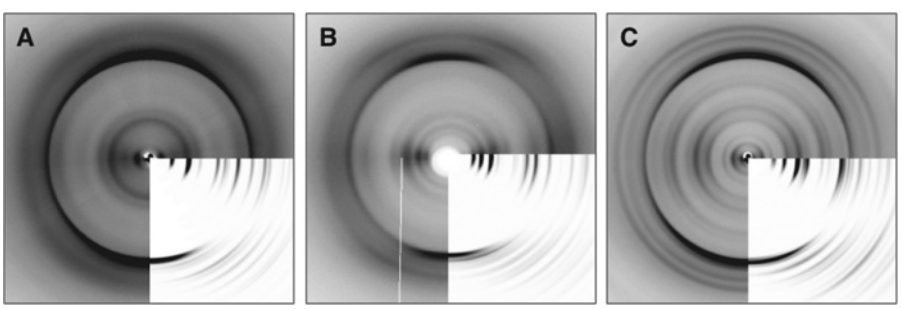

Figure 6 Experimental XRFD patterns compared with patterns simulated from modelled structures

XRFD patterns are shown for alignments of (**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI. Fibre axes are vertical and insets for each show the simulated diffraction pattern for comparison with the experimental data. See also Supplementary Tables S3–S5 (at http://www.biochemj.org/bj/450/bj4500275add.htm). RVFNIM is compared with a film-textured alignment, see Supplementary Figure S5 (at http://www.biochemj.org/bj/450/bj4500275add.htm).

© The Authors Journal compilation © 2013 Biochemical Society

to one another. RVFNIM and VIYKI adopt more classical steric zippers in classes IV and II respectively; the former has parallel β-sheets in a face-to-back and up-down arrangement, whereas the latter β-sheets are face-to-back in an up-up arrangement. The meridional reflections of RVFNIM and VIYKI were better reproduced with β-strands displaced along the fibre axis by $c/2$.

DISCUSSION

The present study confirms further the Waltz algorithm's ability to predict ordered aggregation, and we have characterized two hexapeptides and one pentapeptide uniquely identified during the iterative sequence exploration that developed Waltz [18]. The sequences are found to have varying secondary structure propensities, but in their native structures adopt a conformation dependent on their surrounding amino acid sequence and structural environment. When removed from their native sequences, the Waltz peptides form a β-strand conformation that adopts a cross-β architecture, but with differing lateral packing arrangements. Previous work by Johansson and co-workers suggested that amyloidogenic sequences were often 'promiscuous' and able to adopt both α-helical and β-strand conformations depending on their context and environment [44]. The results of the present study re-enforce this and the concept that the flanking amino acid sequences in amyloidogenic portions of proteins protect against aggregation and self-assembly by forcing the adoption of particular secondary and tertiary structures [45].

On the nanoscopic scale these systems are observed to adopt characteristic and discrete fibril morphologies that may be based on the underlying differences in molecular packing, but are ultimately typical for amyloid fibrils [46]. The exact relationship between these two levels of structure remains unclear. To model the molecular packing of these peptides, the investigations reported in the present study have required careful consideration of the use of the biophysical and structural techniques employed. It was found that CD data is highly dependent on the fibrillar nature of the assemblies, but we highlight methods, as reported elsewhere [22,27], for the successful interpretation of this data. The identification of LD artefacts in CD data has been reported previously [47] and the results of the present study clearly reiterate that LD artefacts can occur in CD experiments, the identification of which have broad methodological implications for using CD to study anisotropic systems. Clearly care should be taken in analyses of these sorts of data, but the previously described phenomena can be usefully rationalized and used in LD experiments to reveal information about chromophore orientation.

Particular care was taken to ensure that the hydrated state of these systems was the same as the dried, through the use of a semi-enclosed chamber to monitor XRFD in real-time over the course of fibril alignment. The structures reported in the present study are found to be unaffected by drying as reported previously [48].

Using XRFD and Clearer, unit cells were determined for the Waltz peptides in the fibrillar state and the architectures that best represent the fibril structure of the Waltz peptides were determined. These peptides do not crystallize in the conditions tested by us (results not shown) and we have focussed on the fibrillar structure due to concerns of the comparability of crystal structures with the absolute structure of self-assembled fibres [40,49]. Despite this, the proposed crystalline class models of amyloid systems provide a useful framework within which to structurally explore the possible architectures of short amyloidogenic sequences. It is probable that the Waltz fibrils contain a mixture of these architectures where the difference between simulated patterns of different class models is slight. It must be considered that more structural information may be obtained from crystal structures, but they represent a single global minimum, which may not be representative of the local minima of the many polymorphs present in real amyloid fibrils.

Although likely to be polymorphic, we have judged quantitatively the most representative predominant polymorphs that constitute the Waltz peptide assemblies. Structurally, RVFNIM and VIYKI are found to adopt arrangements with displacement of β-sheets along the fibre axis ($c/2$) with respect to one another. This has been observed for crystallized short amyloidogenic peptides [12], but, at the time of writing, the results from the present study are the first example of this being demonstrated through experimental and simulated XRFD. The arrangements adopted are variable between the Waltz peptides, which is perhaps surprising given the position sequence similarity of these peptides, but is a reflection of the adaptability of this conformation. Interestingly, aromatic amino acids appear to provide a possible driving force for assembly in some amyloidogenic peptides [20,40] and others have suggested an important contribution of hydrophobicity [50]. Examination of the peptide arrangements reveals that, although hydrophobicity may drive elongation of β-sheets, it is likely that charge and aromaticity modulate the possible arrangements and architecture within the final fibrillar structure via sheet–sheet interactions and lateral association of protofilaments.

In the future, studies of these systems have the potential to make valuable and unique contributions to the model systems currently describing the structural space of amyloid assemblies. The characterization of these systems not only further confirms the ability of the Waltz algorithms to predict ordered amyloid aggregation, but also introduces an opportunity to further understand this conformation while presenting new highly ordered and well-characterized systems for nanotechnological applications.

AUTHOR CONTRIBUTION

Kyle Morris conducted the experimental work, performed analysis and wrote the paper. Alison Rodger provided equipment and expertise and contributed to data analysis. Matthew Hicks contributed to LD data collection and analysis. Maya Debulpaep collected FTIR data. Joost Schymkowitz and Frederic Rousseau contributed materials and wrote the Waltz algorithm on which the paper is based. They also advised on the interpretation and the paper. Louise Serpell managed the project, analysed the results and wrote the paper.

ACKNOWLEDGEMENTS

XRFD data collected at the Diamond I24 microfocus beamline was done so with the expert help of Dr Gwyndaf Evans, Dr Danny Axford and Dr Robin Owen. We thank Dr Julian Thorpe for essential electron microscopy support, Dr Peter Varnai for valuable advice on model minimization and Youssra Al-Hilaly for insightful discussions. The apparatus used to generate the real-time XRFD data was based on a prototype design kindly lent to us by Professor Pawel Sikorski, to whom we are indebted.

FUNDING

This work was supported by the Biotechnology and Biological Sciences Research Council funded Synthetic components network awarded to L.C.S. and K.L.M. L.C.S. is supported by Alzheimer's Research UK. The Switch Laboratory was supported by grants from the Flanders Institute for Biotechnology (VIB), The University of Leuven, the Funds for Scientific Research Flanders (FWO), the Flanders Institute for Science and Technology (IWT) and the Federal Office for Scientific Affairs, Belgium IUAP P7.

REFERENCES

1 Morris, K. and Serpell, L. (2010) From natural to designer self-assembling biopolymers, the structural characterisation of fibrous proteins & peptides using fibre diffraction. Chem. Soc. Rev. **39**, 3445–3453

© The Authors Journal compilation © 2013 Biochemical Society

2 Dobson, C. M. (2001) The structural basis of protein folding and its links with human disease. Philos. Trans. R. Soc. London Ser. B **356**, 133–145
3 Adler-Abramovich, L., Vaks, L., Carny, O., Trudler, D., Magno, A., Caflisch, A., Frenkel, D. and Gazit, E. (2012) Phenylalanine assembly into toxic fibrils suggests amyloid etiology in phenylketonuria. Nat. Chem. Biol. **8**, 701–706
4 Luhrs, T., Ritter, C., Adrian, M., Riek-Loher, D., Bohrmann, B., Dobeli, H., Schubert, D. and Riek, R. (2005) 3D structure of Alzheimer's amyloid-β(1–42) fibrils. Proc. Natl. Acad. Sci. U.S.A. **102**, 17342–17347
5 Wille, H., Bian, W., McDonald, M., Kendall, A., Colby, D. W., Bloch, L., Ollesch, J., Borovinskiy, A. L., Cohen, F. E., Prusiner, S. B. and Stubbs, G. (2009) Natural and synthetic prion structure from X-ray fiber diffraction. Proc. Natl. Acad. Sci. U.S.A. **106**, 16990–16995
6 Petkova, A. T., Ishii, Y., Balbach, J. J., Antzutkin, O. N., Leapman, R. D., Delaglio, F. and Tycko, R. (2002) A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. Proc. Natl. Acad. Sci. U.S.A. **99**, 16742–16747
7 Madine, J., Copland, A., Serpell, L. C. and Middleton, D. A. (2009) Cross-β spine architecture of fibrils formed by the amyloidogenic segment NFGSVQFV of medin from solid-state NMR and X-ray fiber diffraction measurements. Biochemistry **48**, 3089–3099
8 Makin, O. S., Atkins, E., Sikorski, P., Johansson, J. and Serpell, L. C. (2005) Molecular basis for amyloid fibril formation and stability. Proc. Natl. Acad. Sci. U.S.A. **102**, 315–320
9 Jahn, T. R., Makin, O. S., Morris, K. L., Marshall, K. E., Tian, P., Sikorski, P. and Serpell, L. C. (2010) The common architecture of cross-β amyloid. J. Mol. Biol. **395**, 717–727
10 Geddes, A. J., Parker, K. D., Atkins, E. D. and Beighton, E. (1968) "Cross-β" conformation in proteins. J. Mol. Biol. **32**, 343–358
11 Kirschner, D. A., Abraham, C. and Selkoe, D. J. (1986) X-ray diffraction from intraneuronal paired helical filaments and extraneuronal amyloid fibers in Alzheimer disease indicates cross-β conformation. Proc. Natl. Acad. Sci. U.S.A. **83**, 503–507
12 Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J., McFarlane, H. T. et al. (2007) Atomic structures of amyloid cross-β spines reveal varied steric zippers. Nature **447**, 453–457
13 Rousseau, F., Schymkowitz, J. and Serrano, L. (2006) Protein aggregation and amyloidosis: confusion of the kinds? Curr. Opin. Struct. Biol. **16**, 118–126
14 Thompson, M. J., Sievers, S. A., Karanicolas, J., Ivanova, M. I., Baker, D. and Eisenberg, D. (2006) The 3D profile method for identifying fibril-forming segments of proteins. Proc. Natl. Acad. Sci. U.S.A. **103**, 4074–4078
15 Zibaee, S., Makin, O. S., Goedert, M. and Serpell, L. C. (2007) A simple algorithm locates β-strands in the amyloid fibril core of α-synuclein, Aβ, and tau using the amino acid sequence alone. Protein Sci. **16**, 906–918
16 Tartaglia, G. G., Pawar, A. P., Campioni, S., Dobson, C. M., Chiti, F. and Vendruscolo, M. (2008) Prediction of aggregation-prone regions in structured proteins. J. Mol. Biol. **380**, 425–436
17 Fernandez-Escamilla, A.-M., Rousseau, F., Schymkowitz, J. and Serrano, L. (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat. Biotechnol. **22**, 1302–1306
18 Maurer-Stroh, S., Debulpaep, M., Kuemmerer, N., de la Paz, M. L., Martins, I. C., Reumers, J., Morris, K. L., Copland, A., Serpell, L., Serrano, L. et al. (2010) Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. Nat. Methods **7**, 237–242
19 de la Paz, M. and Serrano, L. (2004) Sequence deteminants of amyloid fibril formation. Proc. Natl. Acad. Sci. U.S.A. **101**, 87–92
20 Marshall, K. E., Morris, K. L., Charlton, D., O'Reilly, N., Lewis, L., Walden, H. and Serpell, L. C. (2011) Hydrophobic, aromatic, and electrostatic interactions play a central role in amyloid fibril formation and stability. Biochemistry **50**, 2061–2071
21 Abramoff, A. M., Magelhaes, P. J. and Ram, S. J. (2004) Image processing with ImageJ. Biophotonics Int. **11**, 36–42
22 Andersen, C. B., Hicks, M. R., Vetri, V., Vandahl, B., Rahbek-Nielsen, H., Thogersen, H., Thogersen, I. B., Enghild, J. J., Serpell, L. C., Rischel, C. and Otzen, D. E. (2010) Glucagon fibril polymorphism reflects differences in protofilament backbone structure. J. Mol. Biol. **397**, 932–946
23 Whitmore, L. and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res. **32**, 668–673
24 Compton, L. A. and Johnson, W. C. (1986) Analysis of protein circular-dichroism spectra for secondary structure using a simple matrix multiplication. Anal. Biochem. **155**, 155–167
25 Lees, J. G., Miles, A. J., Wien, F. and Wallace, B. A. (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. Bioinformatics **22**, 1955–1962
26 Mao, D., Wachter, E. and Wallace, B. A. (1982) Folding of the mitochondrial proton adenosine-triphosphatase proteolipid channel in phospholipid-vesicles. Biochemistry **21**, 4960–4968
27 Dafforn, T. R., Rajendra, J., Halsall, D. J., Serpell, L. C. and Rodger, A. (2004) Protein fiber linear dichroism for structure determination and kinetics in a low-volume, low-wavelength couette flow cell. Biophys. J. **86**, 404–410
28 Morris, K. L. and Serpell, L. C. (2012) X-ray fibre diffraction studies of amyloid fibrils. In Amyloid Proteins: Methods and Protocols, Second Edition (Sigurdsson, E. M., Calero, M. and Gasset, M., eds), pp. 121–135, Springer
29 McDonald, M., Kendalla, A., Tanaka, M., Weissman, J. S. and Stubbs, G. (2008) Enclosed chambers for humidity control and sample containment in fiber diffraction. J. Appl. Crystallogr. **41**, 206–209
30 Makin, O. S., Sikorski, P. and Serpell, L. C. (2007) CLEARER: a new tool for the analysis of X-ray fibre diffraction patterns and diffraction simulation from atomic structural models. J. Appl. Crystallogr. **40**, 966–972
31 Wang, J. M., Cieplak, P. and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J. Comput. Chem. **21**, 1049–1074
32 Reference deleted
33 Brooks, B. R., Brooks, III, C. L., Mackerell, Jr, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S. et al. (2009) CHARMM: the biomolecular simulation program. J. Comput. Chem. **30**, 1545–1614
34 Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L. and Schulten, K. (2005) Scalable molecular dynamics with NAMD. J. Comput. Chem. **26**, 1781–1802
35 Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. J. Mol. Graphics Model. **14**, 33–38
36 Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. **292**, 195–202
37 Barth, A. and Zscherp, C. (2002) What vibrations tell us about proteins. Q. Rev. Biophys. **35**, 369–430
38 Miyazawa, T. and Blout, E. R. (1961) Infrared spectra of polypeptides in various conformations – amide I and II bands. J. Am. Chem. Soc. **83**, 712–719
39 Rodger, A., Marrington, R., Geeves, M. A., Hicks, M., de Alwis, L., Halsall, D. J. and Dafforn, T. R. (2006) Looking at long molecules in solution: what happens when they are subjected to Couette flow? Phys. Chem. Chem. Phys. **8**, 3161–3171
40 Marshall, K. E., Hicks, M. R., Williams, T. L., Hoffmann, S. V., Rodger, A., Dafforn, T. R. and Serpell, L. C. (2010) Characterizing the assembly of the Sup35 yeast prion fragment, GNNQQNY: structural changes accompany a fiber-to-crystal switch. Biophys. J. **98**, 330–338
41 Adachi, R., Yamaguchi, K.-i., Yagi, H., Sakurai, K., Naiki, H. and Goto, Y. (2007) Flow-induced alignment of amyloid protofilaments revealed by linear dichroism. J. Biol. Chem. **282**, 8978–8983
42 Hicks, M. R., Kowalski, J. and Rodger, A. (2010) LD spectroscopy of natural and synthetic biomaterials. Chem. Soc. Rev. **39**, 3380–3393
43 Rodger, A. and Norden, B. (1997) Circular Chroism & Linear Dichroism, Oxford University Press, Oxford
44 Kallberg, Y., Gustafsson, M., Persson, B., Thyberg, J. and Johansson, J. (2001) Prediction of amyloid fibril-forming proteins. J. Biol. Chem. **276**, 12945–12950
45 Reumers, J., Maurer-Stroh, S., Schymkowitz, J. and Rousseau, F. (2009) Protein sequences encode safeguards against aggregation. Hum. Mutat. **30**, 431–437
46 Goldsbury, C. S., Cooper, G. J. S., Goldie, K. N., Muller, S. A., Saafi, E. L., Gruijters, W. T. M. and Misur, M. P. (1997) Polymorphic fibrillar assembly of human amylin. J. Struct. Biol. **119**, 17–27
47 Norden, B. (1977) Linear and circular dichroism of polymeric pseudoisocyanine. J. Phys. Chem. **81**, 151–159
48 Squires, A. M., Devlin, G. L., Gras, S. L., Tickler, A. K., MacPhee, C. E. and Dobson, C. M. (2006) X-ray scattering study of the effect of hydration on the cross-β structure of amyloid fibrils. J. Am. Chem. Soc. **128**, 11738–11739
49 Lewandowski, J. R., van der Wel, P. C. A., Rigney, M., Grigorieff, N. and Griffin, R. G. (2011) Structural complexity of a composite amyloid fibril. J. Am. Chem. Soc. **133**, 14686–14698
50 Tracz, S. M., Abedini, A., Driscoll, M. and Raleigh, D. P. (2004) Role of aromatic interactions in amyloid formation by peptides derived from human amylin. Biochemistry **43**, 15901–15908

Received 23 November 2012/17 December 2012; accepted 20 December 2012
Published as BJ Immediate Publication 20 December 2012, doi:10.1042/BJ20121773

© The Authors Journal compilation © 2013 Biochemical Society

Biochem. J. (2013) **450**, 275–283 (Printed in Great Britain) doi:10.1042/BJ20121773

online data

SUPPLEMENTARY ONLINE DATA

Exploring the sequence–structure relationship for amyloid peptides

Kyle L. MORRIS*, Alison RODGER†, Matthew R. HICKS‡, Maya DEBULPAEP§, Joost SCHYMKOWITZ‖, Frederic ROUSSEAU‖ and Louise C. SERPELL*

*School of Life Sciences, University of Sussex, Falmer, East Sussex BN1 9QG, U.K., †Department of Chemistry, University of Warwick, Coventry CV4 7AL, U.K., ‡School of Biosciences, University of Birmingham, Birmingham B15 2TT, U.K., §VIB Switch Laboratory, Free University of Brussels, Pleinlaan 2, Brussels 1050, Belgium, and ‖VIB Switch Laboratory, Department of Cellular and Molecular Medicine, University of Leuven, Leuven 3000, Belgium

HYFNIF

On the basis of the dimensions of a β-strand monomer of HYFNIF, the available packing arrangements within the determined unit cell (Table S4) are limited and not described by the steric zipper classes. The indexed unit cell may only contain two peptides end on end to satisfy the β-sheet spacing (on the basis of equatorial signal measurement) and appropriately accommodate the length of the peptide. A low-resolution reflection at ~40 Å (Table S3) indicates a longer-range repeat, that we hypothesize arises from a protofilament width. The repeating cell of HYFNIF is well described as two β-strands end on end. Model building explored the possibility of the two β-sheets being related by a displacement along the fibre axis, a rotational symmetry or translation symmetry.

RVFNIM

Two monomers of RVFNIM may be accommodated within the determined unit cell (Table S4). The *a* dimension corresponds approximately to the peptide length and the *b* dimension approximately to double the sheet spacing distance of 10.8 Å. In this respect, this systems arrangement was sufficiently explored by the use of the amyloidogenic class models [6].

VIYKI

Four monomers of VIYKI may be arranged within the repeating cell (Table S4). The *a* dimension corresponds to approximately two peptide β-stands of VIYKI and when considering the β-sheet spacing of 9.21 Å, the *b* dimension can accommodate approximately two β-sheets. This creates a large cell that contains four VIYKI monomers. We consider that this is akin to two amyloid class models side by side and so modelling was based on searching the structure space based on a cell containing two β-strands. The *a* dimension may again correspond to a protofilament width and arise from lateral packing of protofilaments.

[1] To whom correspondence should be addressed (email l.c.serpell@sussex.ac.uk).

© The Authors Journal compilation © 2013 Biochemical Society

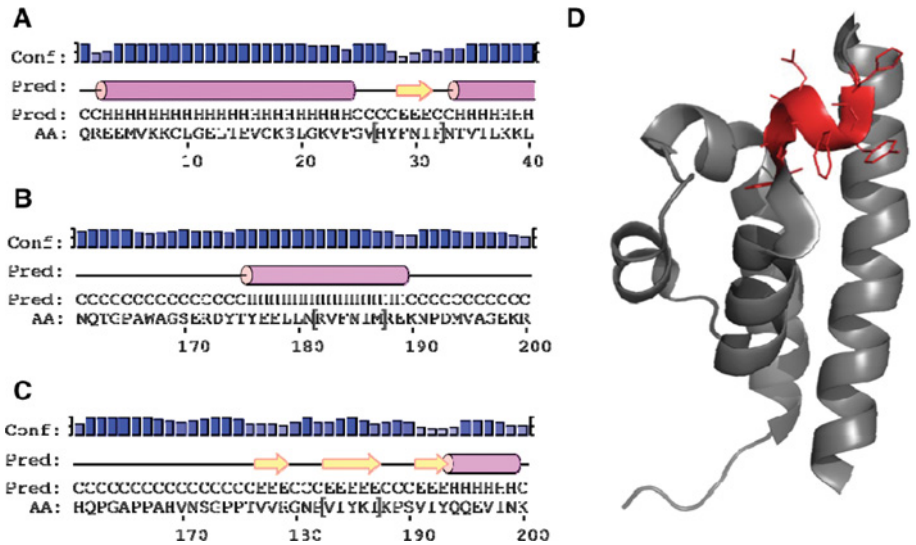

Figure S1 The native sequence context and secondary structures predicted for the Waltz peptides

(**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI in their native full-length sequences. (**D**) The three-dimensional structure of HYFNIF (highlighted in red) in its native structure is shown from PDB code 2KV2 [3]. Secondary structure predictions were made using the online server PSIPred [4,5]. See Table 1 of the main text.

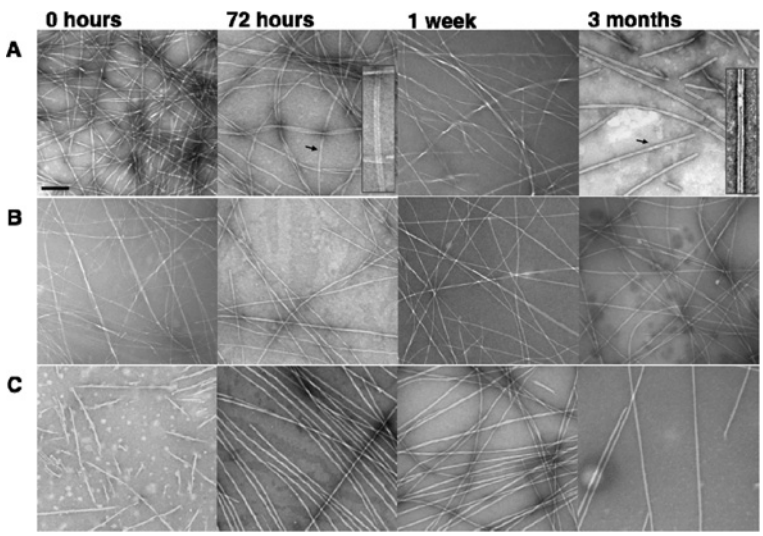

Figure S2 Transmission electron micrographs of the Waltz peptides over time

(**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI are shown at 0 and 72 h to 1 week and 3 months after dissolution. Scale bar, 200 nm. See Figure 1 of the main text.

© The Authors Journal compilation © 2013 Biochemical Society

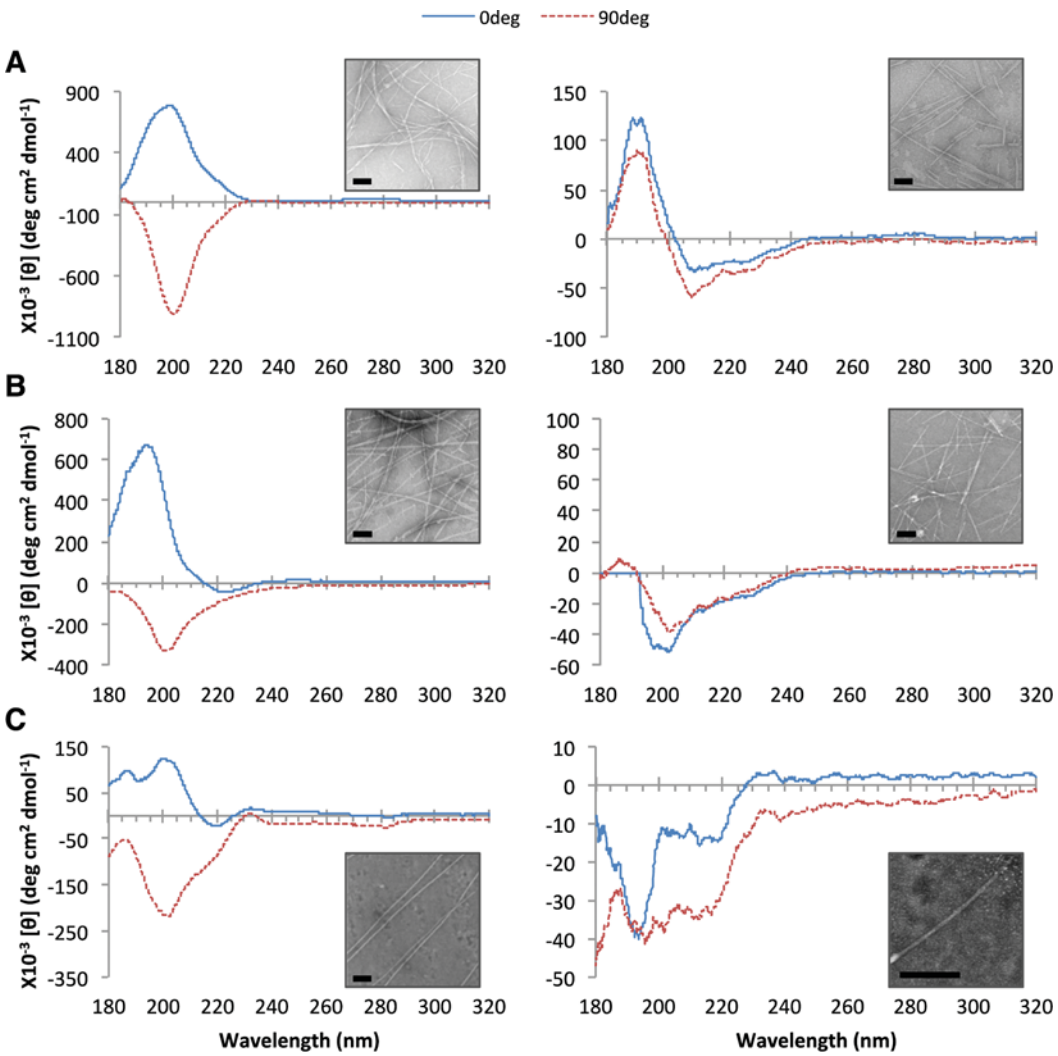

Figure S3 CD spectra for the Waltz peptides pre- and post-sonication

(**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI pre- (left-hand panel) and post- (right-hand panel) bath sonication. The legend indicates sample orientation in the instrument. TEM of the samples are shown as insets; scale bars, 100 nm. LD artefacts are observed in the samples pre-sonication as indicated by sample orientation signal dependence. The true CD signals reveal β-sheet secondary structure with additional contributions from phenylalanine and tyrosine at ~210 nm and ~230 nm respectively, due to the high aromatic content of the Waltz peptides (17–33 %). See Table S1 and Figure 2 of the main text.

© The Authors Journal compilation © 2013 Biochemical Society

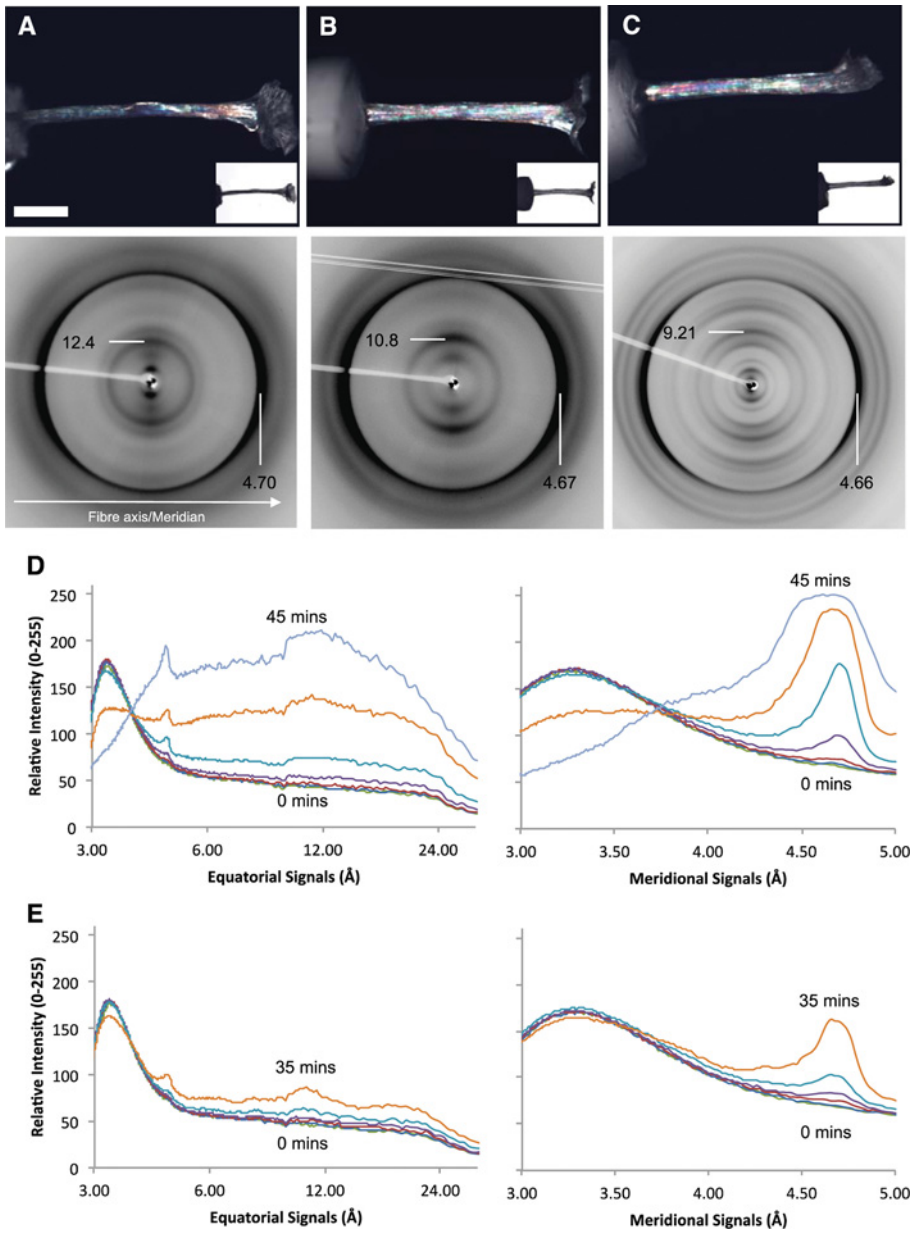

Figure S4 Typical alignments, fibrous and real-time XRFD of the Waltz peptides

(**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI as visualized by cross-polarized microscopy. The insets show the alignments under normal light microscopy. Scale bar, 500 μm. XRFD from the corresponding alignments is shown with major meridional (vertical lines) and equatorial (horizontal lines) reflections labelled in Å. Data were collected at Diamond I24. Real-time XRFD from (**D**) HYFNIF and (**E**) RVFNIM is shown graphically. In both cases the equatorial and meridional signals characteristic of the amyloid structure are observed, amd the diffuse diffraction signal at $\sim$3 Å indicates the alignment is still hydrated. Equatorial signals are shown over a logarithmic scale for clarity. See Figure 4 in the main text.

© The Authors Journal compilation © 2013 Biochemical Society

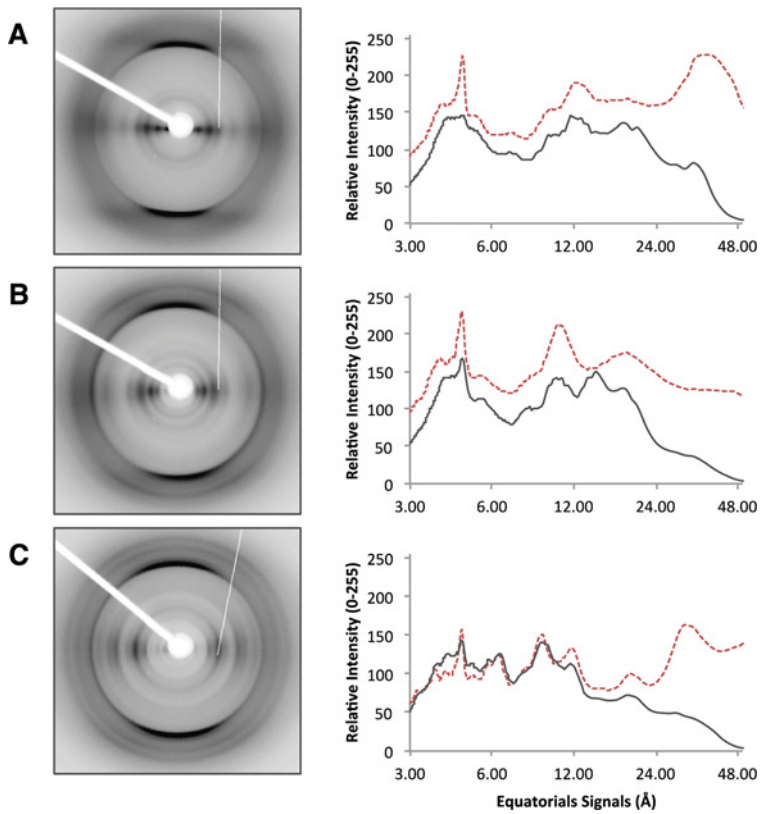

Figure S5 XRFD exhibited from different textured alignments of the Waltz peptides

(**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI in a film-texture. Graphical traces of the equators are shown for comparison of (black line) film-textured to (red dotted line) fibrous-textured alignments. These indicate the same structure is reported with some additional information in the case of RVFNIM, which is later used for comparison with simulation data. Patterns are aligned with equivalent fibre axes vertical and equatorial signals are shown over a logarithmic scale for clarity. See Figure 6 of the main text.

© The Authors Journal compilation © 2013 Biochemical Society

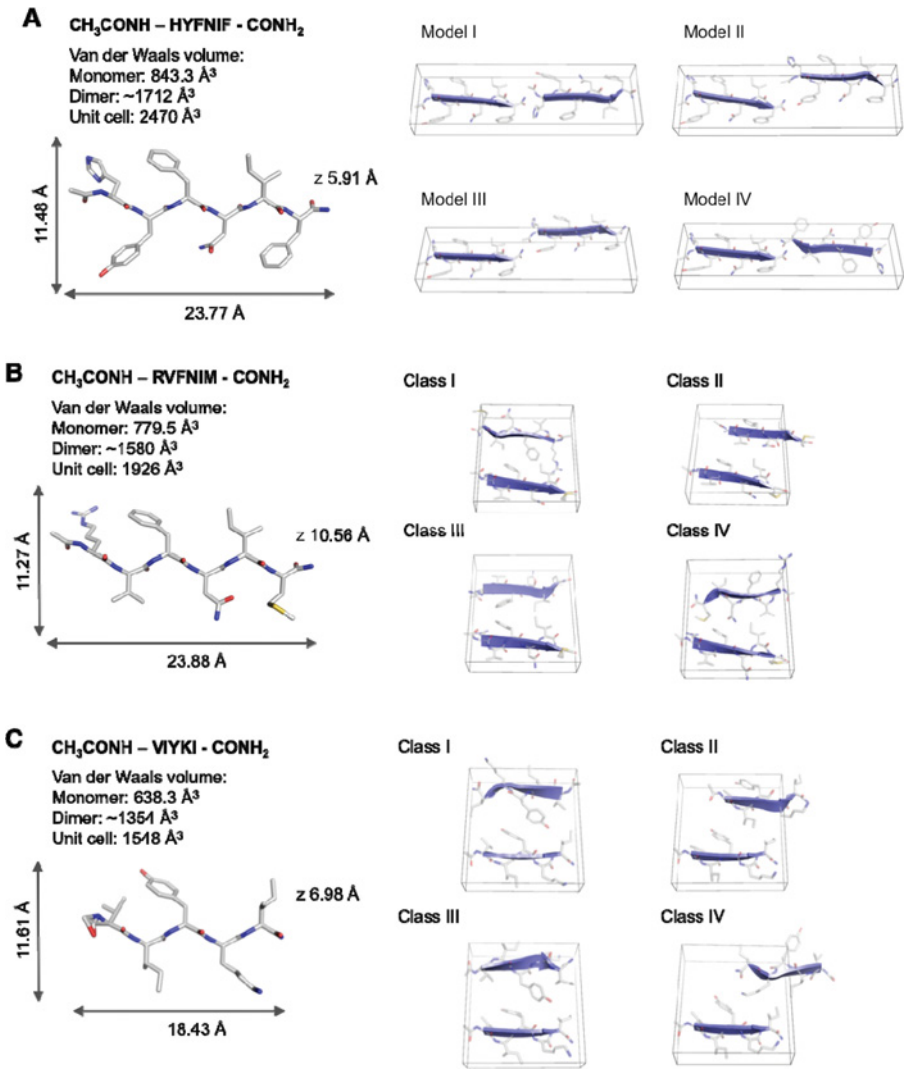

Figure S6 Four structural classes constructed for the three Waltz peptides

(**A**) HYFNIF, (**B**) RVFNIM and (**C**) VIYKI. See Figure 5 of the main text.

© The Authors Journal compilation © 2013 Biochemical Society

Table S1 Dichroweb analysis of the CD spectra recorded for the Waltz peptides post-sonication, highlighting the predominant secondary structural elements

Analysis was performed over two data sets using the CDSSTR analysis program [1] and SP175 190–240 nm reference set [2]. NRMSD, normalized root-mean-square deviation.

	HYFNIF		RVFNIM		VIYKI	
Secondary structure	%	±	%	±	%	±
Helix	6.0	1.4	4.5	2.1	2.0	0.0
Strand	42.5	3.5	40.0	5.7	42.5	0.7
Turn	11.5	0.7	12.5	2.1	12.0	0.0
Other	40.0	1.4	42.5	0.7	41.5	0.7
Total	100.0	7.1	99.5	10.6	98.0	1.4
NRMSD	0.03	0.01	0.05	0.00	0.04	0.01

Table S2 LD peaks exhibited by the Waltz fibrils when under shear and Couette flow alignment

LD and LD_{red} values are reported in $\Delta Abs \times 10^4$ after background subtraction and zeroing. Abs_{iso} are reported as the isotropic absorbance from the sample after background Abs_{iso} subtraction. Shear aligned fibrils are perpendicular to the orientation axis (⊥), whereas flow aligned fibrils are parallel with the orientation axis (//).

	HYFNIF				RVFNIM				VIYKI			
	nm	LD	Abs_{iso}	LD_{red}	nm	LD	Abs_{iso}	LD_{red}	nm	LD	Abs_{iso}	LD_{red}
Shear aligned	283.4	− 3.862	0.003	− 1492	–	–	–	–	283.0	− 8.174	0.017	− 492.3
	278.0	− 4.278	0.004	− 1033	–	–	–	–	276.0	− 7.777	0.020	− 382.2
	234.0	1.881	0.027	69.63	–	–	–	–	231.2	12.16	0.083	146.5
	200.0	− 198.4	0.490	− 404.8	196.4	− 268.7	0.429	− 626.7	200.4	− 200.1	0.628	− 318.7
Flow aligned	282.6	18.68	0.004	4861	–	–	–	–	283.0	4.655	0.011	429.1
	277.2	19.61	0.005	3794	–	–	–	–	276.0	3.878	0.014	272.0
	234.4	− 10.53	0.026	− 399.6	–	–	–	–	231.6	− 8.625	0.064	− 135.1
	200.4	964.4	0.491	1963	–	–	–	–	200.0	98.65	0.525	187.9

Table S3 XRFD signals exhibited from fibrous alignments of HYFNIF, RVFNIM and VIYKI

Reflections measured manually are indicated by *. E, equatorial; M, meridional; Rad, radially averaged.

HYFNIF			RVFNIM			VIYKI		
Signal (Å)	Normalized relative intensity	Axis	Signal (Å)	Normalized relative intensity	Axis	Signal (Å)	Normalized relative intensity	Axis
38.98	0.95	E				33.26*	0.47	E
19.22	0.70	E	18.72	0.73	E	20.04	0.29	E
12.37	0.78	E	10.79	0.87	E	11.81	0.38	E
10.03	0.63	E	8.31*	0.62	E	9.21	0.43	E
7.05	0.50	E	5.48	0.59	E	8.09*	0.34	E
5.23	0.60	E				6.35	0.36	E
						5.82*	0.34	E
						5.12	0.28	E
						3.25*	0.22	E
4.70	1.00	M	4.67	1.00	M	4.66	1.00	M
3.95	0.66	Rad	2.38	0.27	M	4.01	0.34	M
2.38	0.20	M				3.72	0.34	Rad
						2.38*	0.10	M

© The Authors Journal compilation © 2013 Biochemical Society

Table S4 The diffraction signals and their respective indexing to the modelled unit cells predicted by Clearer

[0 0 1] indices correspond to the fibre axis, all others are perpendicular to the fibre axis. The unit cell dimensions correspond to the peptides as follows: $a = \beta$-strand long axis, $b = \beta$-sheet spacing and $c = \beta$-sheet hydrogen bonding distance/fibre axis. The intensity of the 011 reflection is increased for RVFNIM and VIYKI due to β-sheet displacement along the fibre axis leading to higher resolution observed major meridional values.

HYFNIF					RVFNIM					VIYKI				
	a (Å) 41.72	*b* (Å) 12.60	*c* (Å) 4.70	$\alpha = \beta = \gamma$ 90°		*a* (Å) 18.80	*b* (Å) 21.94	*c* (Å) 4.78	$\alpha = \beta = \gamma$ 90°		*a* (Å) 34.94	*b* (Å) 19.01	*c* (Å) 4.76	$\alpha = \beta = \gamma$ 90°
Indexing														
Obs	h	k	l	Calc	Obs	h	k	l	Calc	Obs	h	k	l	Calc
38.98	1	0	0	41.72	18.72	1	0	0	18.80	33.26	1	0	0	34.95
19.22	2	0	0	20.86	10.79	0	2	0	10.97	20.04	0	1	0	19.01
12.37	0	1	0	12.60	8.31	2	1	0	8.64	11.81	3	0	0	11.65
10.03	4	0	0	10.43	5.48	0	4	0	5.49	9.21	1	2	0	9.17
7.05	6	0	0	6.95	4.67	0	0	1	4.78	8.09	4	1	0	7.94
5.23	8	0	0	5.22	4.67	0	1	1	4.67	6.35	0	3	0	6.34
4.70	0	0	1	4.70	2.38	0	0	2	2.39	5.82	6	0	0	5.82
4.70	1	0	1	4.67						5.12	4	3	0	5.13
3.95	6	0	1	3.89						3.25	8	4	0	3.22
2.38	0	0	2	2.35						4.66	0	0	1	4.76
										4.66	0	1	1	4.62
										4.01	2	2	1	4.14
										2.38	0	0	2	2.38

Table S5 Quantitative RF comparisons of the simulated with the experimental fibre diffraction data

HYFNIF	RF	RVFNIM	RF	VIYKI	RF
Model I	0.37	Class I	0.33	Class I	0.41
Model II	0.38	Class II	0.35	Class II	0.33
Model III	0.34	Class III	0.39	Class III	0.50
Model IV	0.35	Class IV	0.32	Class IV	0.47

REFERENCES

1 Compton, L. A. and Johnson, W. C. (1986) Analysis of protein circular-dichroism spectra for secondary structure using a simple matrix multiplication. Anal. Biochem. **155**, 155–167

2 Lees, J. G., Miles, A. J., Wien, F. and Wallace, B. A. (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. Bioinformatics **22**, 1955–1962

3 Kim, Y. M. and Choi, B.-S. (2010) Structure and function of the regulatory HRDC domain from human Bloom syndrome protein. Nucleic Acids Res. **38**, 7764–7777

4 Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. **292**, 195–202

5 Buchan, D. W. A., Ward, S. M., Lobley, A. E., Nugent, T. C. O., Bryson, K. and Jones, D. T. (2010) Protein annotation and modelling servers at University College London. Nucleic Acids Res. **38**, 563–568

6 Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J., McFarlane, H. T. et al. (2007) Atomic structures of amyloid cross-β spines reveal varied steric zippers. Nature **447**, 453–457

Received 23 November 2012/17 December 2012; accepted 20 December 2012
Published as BJ Immediate Publication 20 December 2012, doi:10.1042/BJ20121773

© The Authors Journal compilation © 2013 Biochemical Society